**ANALISIS REGRESI SEDERHANA**

 Analisis regresi pada umumnya digunakan untuk keperluan uji hipotesis yang berkaitan dengan prediksi dan eksplanasi. Contoh pemakaian analisis regresi untuk keperluan prediksi adalah penggunaan skor tes SNMPTN untuk memprediksikan/meramalkan keberhasilan belajar calon mahasiswa di perguruan tinggi. Sedangkan pemakaian analisis regresi untuk keperluan eksplanasi adalah untuk mengungkapkan faktor-faktor yang menentukan suatu fenomena. Misalnya, untuk mengungkapkan faktor-faktor yang mempengaruhi prestasi belajar seseorang. Permasalahan yang pertama tentang prediksi dan permasalahan yang kedua tentang eksplanasi, keduanya bisa diselesaikan dengan teknik analisis regresi. Pada dasarnya teknik ini berusaha menganalisis peranan setiap ubahan (prediktor) dalam menentukan besarnya varians dari suatu ubahan yang ada pada fenomena yang diteliti (ubahan kriterium).

 Seperti pada teknik korelasi, teknik analisis regresi menggunakan asumsi adanya hubungan yang linier atau berupa garis lurus antara variabel prediktor dengan variabel kriterium. Teknik analisis regresi yang menggunakan asumsi hubungan linier disebut regresi linier. Adapun bentuk umum persamaan garis regresi Y atas X, dimana X adalah ubahan prediktor adalah :

 Y’ = a + b X → regresi sederhana dengan sebuah ubahan X

Y’ = a + b1.X1 + b2.X2 + ………….+bn.Xn

 Pada prinsipnya teknik analisis regresi ini adalah menentukan suatu garis lurus (garis regresi/garis prediksi) yang paling tepat diantara penyebaran titik-titik ploting (scatter-plot). Garis regresi yang tepat ini harus memenuhi syarat sebagai berikut:

1. Berupa garis lurus atau memenuhi persamaan linier
2. Jumlah kuadrat penyimpangan skor pengamatan dan skor yang diramalkan adalah minimum »= minimum (metode least square).

Berdasarkan konsep kuadrat terkecil (least square) inilah, maka koefisien a dan b dapat dihitung.

Rumus :

 → b=

**Contoh :**

 Akan diteliti mengenai peranan motivasi (X) terhadap prestasi belajar (Y) pada sekelompok siswa yang terdiri dari 30 orang. Dari tiap-tiap siswa yang diteliti tersebut dicatat tentang dua karakteristik, yaitu motivasi (X) dan prestasi belajarnya (Y). Adapun datanya disajikan dalam tabel berikut.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Resp.** | **X** | **Y** |  |  | **XY** |
| 1 | 34 | 32 | 1156 | 1024 | 1088 |
| 2 | 38 | 36 | 1444 | 1296 | 1368 |
| 3 | 34 | 31 | 1156 | 961 | 1054 |
| 4 | 40 | 38 | 1600 | 1444 | 1520 |
| 5 | 30 | 29 |  |  |  |
| 6 | 40 | 35 | dst | dst | dst |
| 7 | 40 | 33 |  |  |  |
| 8 | 34 | 30 |  |  |  |
| 9 | 35 | 32 |  |  |  |
| 10 | 39 | 36 |  |  |  |
| 11 | 33 | 31 |  |  |  |
| 12 | 32 | 31 |  |  |  |
| 13 | 42 | 36 |  |  |  |
| 14 | 40 | 37 |  |  |  |
| 15 | 42 | 35 |  |  |  |
| 16 | 42 | 38 |  |  |  |
| 17 | 41 | 37 |  |  |  |
| 18 | 32 | 30 |  |  |  |
| 19 | 34 | 30 |  |  |  |
| 20 | 36 | 30 |  |  |  |
| 21 | 37 | 33 |  |  |  |
| 22 | 36 | 32 |  |  |  |
| 23 | 37 | 34 |  |  |  |
| 24 | 39 | 35 |  |  |  |
| 25 | 40 | 36 |  |  |  |
| 26 | 33 | 32 |  |  |  |
| 27 | 34 | 32 |  |  |  |
| 28 | 36 | 34 |  |  |  |
| 29 | 37 | 32 |  |  |  |
| 30 | 38 | 34 |  |  |  |
| **n= 30** | **1105** | **1001** | **41.029** | **33.599** | **37.094** |

Berdasarkan data-data diatas, maka koefisien a dan b dapat dihitung sebagai berikut:

 atau b=

 = **0,68**

Dengan demikian, persamaan garis regresinya adalah :

Dalam hal ini, Y diganti (Y topi) ⇒ untuk membedakan antara Y ramalan (yang bentuk dasarnya : Y = α +β.X) dengan Y hasil pengamatan (.

**UJI KELINIERAN DAN KEBERARTIAN REGRESI**

 Sebelum digunakan untuk membuat kesimpulan, maka regresi yang diperoleh harus diuji dulu yang berkaitan dengan : (1) kelinieran bentuk regresi (uji linieritas hubungan) ; dan (2) keberartian regresi (koefisien regresi). Pemeriksaan kelinieran bentuk regresi dilakukan melalui pengujian hipotesis bahwa : regresi linier melawan hipotesis tandingan. bahwa : regresi non linier.

 Sementara itu, uji keberartian regresi diperiksa melalui pengujian hipotesis bahwa koefisien-koefisien regresi atau koefisien arh b samadengan nol (tidak berarti) melawan hipotesis tandingan bahwa koefisien arah regresi atau b tidak sama dengan nol.

H0 : Pop. b = 0 → H0 : β = 0 → tidak berarti

H0 : Pop. b ≠ 0 → H0 : β ≠ 0 → berarti

 Untuk uji kelinieran diperlukan adanya beberapa pengulangan atau kelompok data x. Tiap kelompok terdiri atas beberapa data x yang berharga sama, sementara harga-harga Y pasangannya tetap seperti pasangan data semula.

 Dengan pengelompokan data x yang berharga sama tersebut, akan diperoleh bahwa pada kelompok pertama, ada n, buah data x yang masing-masing harganya sama yaitu x1 berpasangan dengan Y yang umumnya harganya berbeda-beda. Pada kelompok dua, ada n2 buah data x yang masing-masing harganya sma yaitu x2 berpasangan dengan Y yang harganya berbeda-beda dan begitu seterusnya. Seluruhnya ada k buah kelompok yang dalam tiap kelompoknya harga-harga x nya sama besar. Berdasarkan data dimuka, maka pengelompokkan harga-harga x yang sama tersebut dapat dilihat pada table berikut:

|  |  |  |  |
| --- | --- | --- | --- |
| **X** | **Kelompok** | **Ni** | **Y** |
| 303232333334343434343536363637373738383939404040404041424242 | 123456789101112 | 122513322513 | 293130313232313030323230323433343236343635383533373637363538 |

 Berdasarkan data pengelompokan harga-harga yang telah disajikan pada table diatas, maka uji kelinieran dan keberartian regresi dapat dilakukan terlebih dahulu menghitung jumlah kuadrat (jk) untuk berbagai sumber variasi. Sumber variasi tersebut adalah total atau jk (T), regresi (a), regresi (bIa), sisa atau resida, tuna cocok dan galat atau kekeliruan (error). Jk untuk berbagai sumber variasi tersebut berturut-turut diberi symbol dengan jk (T), jk (a), jk (bIa), jk (s) atau jk res, jk (TC) dan jk (G) atau jk (E), dan dapat dihitung dengan rumus-rumus sebagai berikut :

Dalam persamaan umum dipersatukan manjadi Jk total ⇒

Jk (s) =

Jk (G) = JK (E) =

Jk (TC) = Jk (s) – Jk (G) =

 Setiap sumber variasi tersebut memiliki derajat kebebasan atau dk, yang besarnya berturut-turut n untuk total, 1 untuk regresi (a), 1 untuk regresi (bIa), (n-2) untukn sisa/ residu, (k-2) untuk tuna cocok dan (n-k) untuk galat/ error. Selanjutnya dapat ditentukan harga rerata jumlah kuadrat atau Rjk untuk masing-masing sumber variasi yakni dengan cara membagi jk oleh dk-nya masing-masing.

 Sementara itu, Rjk (bIa) sering pula dilambangkan dengan reg ( Varians regresi), dan Rjk (s) atau Rjk (res) sering dilambanngkan oleh sis atau res. Demikian pula Rjk (TC) juga dapat dilambangkan dengan TC dan Rjk (G) atau Rjk (E) sering dilambangkan dengan G atau E. Semua besaran yang diperoleh selanjutnya disusun dalam daftar analisis varians (ANAVA) untuk regresi linier sederhana (Tabel Rangkuman Regresi).

 Besaran-besaran yang ada dalam dftar anava tersebut, khususnya pada kolom Rjk, digunakan untuk menguji hipotesis :

1. Koefisien arah regresi tidak berarti melawan hipotesis tandingan bahwa koefisien arah regresi berarti H0 : Pop bi = 0 VS Ha : Pop bi.
2. Bentuk regresi linier, melawan hipotesis tandingan bahwa bentuk regresi non linier.

Kedua hipotesis diatas diuji dengan menggunakan statistic F yang dibentuk oleh perbandingan dua Rjk.

 Untuk menguji hipotesis (1) dipakai statistic F= reg / res dan selanjutnya digunakan table distribusi F dengan dk pembilang 1 dan dk penyebut (n-2).

**Kriteria:**

 Tolak hipotesis bahwa koefisien arah regresi tidak berarti (H0:Pop b =0) jika Fh = reg/ res ≥ .

Untuk menguji hipotesis (z) dipakai statistik F= TC / E yang selanjutnya dignakan tabel distribusi F dengan dk pembilang (k-2) dan dk penyebut (n-k).

Kriteria:

Tolak hipotesis bahwa bentuk regresi linier, jika F =TC / E ≥ F (1-α) (k-2 ; n-k).

Dalam hal lainya, hipotesis-hipotesis yang disebutkan diatas diterima.

**Tabel Anava Regresi linier sederhana**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sumber Variasi** | **dk** | **Jk** | **Rjk** | **F** |
| Total | n |  |  | - |
| Regresi (a) | 1 | Jk (a) | Jk (a) |  |
| Regresi (bIa) | 1 | Jk (bIa) |  |  |
| Residu / sisa | n-2 | Jk res |  | - |
| Tuna cocok | k-2 | Jk (TC) |  |  |
| Galat /Error | n-k | Jk (E) |  |  |

Berdasarkan rumus-rumus diatas, maka berdasarkan data yang telah disajikan dimuka selanjutnya akan diuji apakah persamaan garis regresi tersebut memiliki bentuk linier dan koefisien arah regresi yang berarti ataukah tidak.

Jk (a) = = 33.400,03

Jk (bIa) = b.

 = 0,68 = 151,75

Jk (G) = Jk (E) =

Berdasarkan pengelompokan harga x yang sama, ternyata ada 12 kelompok (k=12) dengan n1=1, n2=2, n3=2,…….,n12=3. Maka Jk (G) atau Jk (E) dapat dihitung sebagai berikut :

Jk (G) = Jk (E) = +

=

Jk (G) = Jk (E) = 37,67

Jk (TC) = - Jk (E) = 47,22 – 37,67 ⇒ Jk (TC) = 9,55

⇒ Selanjutnya harga-harga tersebut diatas dimasukkan dalam tabel anava sebagai berikut :

Anava untuk regresi linier

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sumber Variasi** | **dk** | **Jk** | **Rjk** | **F** |
| Total | 30 | 33,599 |  |  |
| Regresi (a)Regresi (bIa)Residu / sisa | 1128 | 33,400,03151,7547,22 | -151,751,69 | -89,79 |
| Tuna cocokGalat /Error | 1018 | 9,5537,67 | 0,962,09 | 0,45 |

Misalkan ditentukan α= 0,05

⇒untuk menguji hipotesis (1), maka harga F tabel =4,20

Ternyata bahwa : F tabel = < Fhitung

Sehingga : H0 : Pop b = 0 ⇒ ditolak

Kesimpulan : Koefisien arah regresi berarti (Fh > F tab).

⇒Untuk menguji hipotesis (2), maka harga F tabel

= 2,43

Ternyata F hitung < ⇒ H0 : diterima

Kesimpulan : Kita terima bahwa bentuk regresi adalah linier.

**KETELITIAN PREDIKSI :**

 Kesalahan prediksi dapat dinyatakan dengan simppangan baku dari Y untuk harga x tertentu. Jadi, regresi yang diperoleh berdasarkan data pengamatan adalah merupakan taksiran terhadap regresi Y=α + βx dalam populasi, dimana n buah pasangan data yang dipakai untuk menentukan telah diambil. Dalam menaksir Y sebenarnya (dalam populasi) berdasarkan regresi akan kita temui suatu penyimpangan (Error). Perbedaan ini merupakan galas (error)cyang apabila dihitung Jk-nya kemudian dibagi dengan dk yaitu (n-2), maka akan diperoleh varians galas (error) taksiran Y untuk x tertentu, yang dilambangkan dengan , dan dapat dihitung dengan rumus sebagai berikut :

 ⇒ Jk residu

 Varians ini merupakan taksiran untuk varians dalam populasi dengan model regresi Y= α + βx.

**Asumsi :**

1. Besarnya varians (Varians Y) untuk tiap kelompok harga x tertentu adalah sama besar.

⇒ Variansnya homogen (disebut : homoscedasticity)

1. Distribusi Y untuk tiap kelompok harga x tertentu adalah normal, dengan rata-rata (α+βx).

⇒ Pengujian dua asumsi ini dapat dilakukan dengan menggunakan scatter plot (scatter diagram).

⇒ Selanjutnya dengan adanya taksiran simpangan baku Y ini, berdasarkan regresi , akan dapat dihitung pula varians untuk konstanta a yaitu dan varians untuk koefisien arah regresi b, yaitu Sb. Adapun rumusnya sebagai berikut :

 adalah rata-rata X untuk semua pengamatan